MPSRCH (TM)

```
*********************************************************************
*********************************************************************
**                                                               **
**                                                               **
**                                                               **
**                                                               **
**                                                               **
**                                                               **
*********************************************************************
*********************************************************************
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 28 20:25:28 1999; MasPar time 603.24 Seconds
                 1002.380 Million cell updates/sec

Tabular output not generated.

Title: >US-09-327-230-1
Description: (1-2822) from US09327230.seq
Perfect Score: 2822
N.A. Sequence: 1 gcaacgcacacagacaagca..............ttccaggtttcggtttcgggg 2822
       Comp: cgttgcgtgtgtctgttccgt..............aaggtccaaaccaaagccc

Scoring table: TABLE default
       Gap 6

Nmatch STD :   Dbase 0; Query 0
Searched: 271905 seqs, 107135622 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database: n-geneseq35

Statistics: Mean 9.803; Variance 6.092; scale 1.609

Pred: No. is the number of results predicted by chance to have a
                 score greater than or equal to the score of the result being printed,
                 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description	Pred. No.
1	2822	100.0	2822	51	V57302	Maize cell death supp	0.00e+00
2	49	1.7	204	1	N81164	Base substituted E.co	5.71e-13
3	44	1.6	91	9	Q51746	Oligonucleotide probe	3.73e-10
4	42	1.5	91	9	Q51746	Oligonucleotide probe	4.74e-09
5	42	1.5	204	1	N81164	Base substituted E.co	4.74e-09
c 6	33	1.2	91	46	V44650	Base substituted E.co	2.2e-04
c 7	33	1.2	91	46	V44650	Mammalian DNA replica	2.82e-04
c 8	34	1.2	114	12	Q70467	Generic DNA sequence	8.68e-05
9	33	1.2	172	32	T76363	Human interleukin 8 a	2.82e-04

10	32	1.1	114	12	Q70465	Generic DNA sequence	9.05e-04
11	32	1.1	114	12	Q70468	Generic DNA sequence	9.05e-04
12	31	1.1	114	12	Q70469	Generic DNA sequence	2.86e-03
13	31	1.1	114	12	Q70470	Generic DNA sequence	2.86e-03
14	31	1.1	114	12	Q70465	Generic DNA sequence	2.86e-03
15	30	1.1	114	12	Q70469	Generic DNA sequence	8.93e-03
16	31	1.1	172	32	T76363	Human interleukin 8 a	2.86e-03
17	30	1.1	178	32	T76405	Human endothelin-1 an	8.93e-03
18	27	1.1	67	24	T14322	Primer used in the la	2.46e-01
19	29	1.0	114	12	Q70472	Generic DNA sequence	2.74e-02
20	29	1.0	114	12	Q70466	Generic DNA sequence	2.74e-02
21	28	1.0	114	12	Q70471	Generic DNA sequence	8.29e-02
c 22	29	1.0	114	12	Q70470	Generic DNA sequence	2.74e-02
c 23	30	1.0	114	12	Q70466	Generic DNA sequence	8.29e-02
c 24	29	1.0	114	12	Q70468	Generic DNA sequence	2.74e-02
c 25	29	1.0	114	12	Q70467	Generic DNA sequence	2.74e-02
26	27	1.0	190	32	T76452	Chymase antisense oli	2.46e-01
27	29	1.0	200	32	T76398	Human leukotriene C4	2.74e-02
28	28	1.0	250	32	Q43519	Substance P antisense	2.46e-01
29	30	1.0	657	7	V57302	Maize cell death supp	2.74e-02
30	31	1.0	2822	51	V57302	Maize cell death supp	2.74e-02
31	32	0.9	67	24	T14322	Degenerate FMN reduct	2.46e-01
32	26	0.9	68	34	T73397	Oligonucleotide tag c	7.18e-01
33	33	0.9	70	24	T63255	Messenger RNA primer	7.18e-01
34	34	0.9	72	20	T14325	Conjugate formed by 1	2.05e+00
35	35	0.9	74	21	T13613	DC43 TSAR library gen	2.05e+00
36	36	0.9	78	59	V82092	DBY746 yeast cell pol	1.1e-01
37	37	0.9	81	21	T13611	DC43 TSAR library gen	2.05e+00
38	38	0.9	81	21	Q70473	Generic DNA sequence	2.05e+00
39	39	0.9	114	12	Q70472	Generic DNA sequence	7.18e-01
40	40	0.9	114	12	Q70473	Generic DNA sequence	7.18e-01
41	41	0.9	114	12	Q70472	Generic DNA sequence	7.18e-01
42	42	0.9	178	32	T76152	Human vascular cell a	7.18e-01
43	43	0.9	390	30	T76405	Human endothelin-1 an	2.05e+00
44	44	0.9	390	30	T67438	H. Pylori secreted or	2.05e+00
c 39		0.9	114	12	Q70472	Generic DNA sequence	7.18e-01
c 40		0.9	114	12	Q70473	Generic DNA sequence	7.18e-01
c 41		0.9	114	12	Q70472	Generic DNA sequence	7.18e-01
c 42		0.9	178	32	T76152	Human vascular cell a	7.18e-01
c 43		0.9	390	30	T67919	H. Pylori secreted or	2.05e+00
c 44		0.9	592	16	085520	DNA probe 26 detects	2.05e+00
```

## ALIGNMENTS

RESULT 1

ID     V57302 standard; DNA; 2822 BP.
AC     V57302;
DT     11-JAN-1999  (first entry)
DE     Maize cell death suppressor gene llsl promoter.
KW     Lethal leaf spot-1; llsl gene; maize; necrosis; cell death;
KW     disease resistance; herbicide resistance; dioxygenase; promoter;
KW     ds.
OS     Zea mays.
PN     WO9839422-A1.
PN     11-SEP-1998;
PD     03-MAR-1998; U04040.
PF     04-MAR-1997; US-610009.
PR     (PION-) PIONEER HI-BRED INT INC.
PA     (UMOR-) UNIV MISSOURI.
PI     Briggs SP, Gray J, Gurmukh JS;
PI     Briggs SP, Gray J;
DR     WPI; 98-506354/43.
PT     New isolated plant cell death suppressing gene - used for e.g.
PT     spread of cell death from wounding or internal stresses that occur
PT     during photosynthesis. The invention relates to methods and
PT     compositions for suppressing cell death and controlling disease
PT     resistance in plants. The compositions, cell death suppressing
PT     engineering herbicide resistance, enhancing transformation efficiency,
PT     engineering herbicide resistance or targeting cell ablations.
PS     Claim 38; Page 51-53; 95pp; English.
CC     This nucleotide sequence comprises the promoter region of the
CC     maize leaf spot-1 (llsl) gene (see V57301 and V57303). The llsl
CC     gene encodes a novel maize protein (see W79001) which inhibits the
CC     spread of cell death from wounding or internal stresses that occur
CC     during photosynthesis. The invention relates to methods and
CC     compositions for suppressing cell death and controlling disease
CC     resistance in plants. The compositions are useful for activating disease
CC     resistance, enhancing cell transformation efficiency, engineering
CC     herbicide resistance and genetically targeting cell ablations.
CC     Novel promoter sequences are provided for the expression of genes
CC     in plants. A chimeric gene comprising the promoter operably
CC     linked with a heterologous coding sequence is claimed, as are a

```
CC   vector comprising the chimeric gene, a host cell comprising the
CC   vector, a plant stably transformed with the chimeric gene, and
CC   transformed seed of the plant.
SQ   Sequence 2822 BP;   851 A;   574 C;   542 G;   855 T;

Query Match        100.0%;  Score 2822;  DB 51;  Length 2822;
Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
Matches 2822;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Db       1 gcaacgcacacaagacaggcagccgaatgtcttcttcgcgggtcagtgaactcactcacacagg    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 ctattcgtcttaagttttttgttcaacatcacatacttgttgctaatgtaacaacaa          60

Db      61 ctattcgtcttaagttttttgttcaacatcacatacttgttgctaatgtaacaacaa        120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61 aaattcacacgcctcacaaacattacaatgattcaaaatataaaataaaacttg          120

Db     121 aaattcacacgcctcacaaacattacaatgattcaaaatataaaataaaacttg           180
Qy     121 gaggacttttgtactgctagagacaacacctactctactgctgcttaccgagacag        180

Db     181 gaggacttttgtactgctagagacaacacctactctactgctgcttaccgagacag         240
Qy     181 aggaaatacacacgaggcaactcgttgttggacttgttgcaaaatagcaaggtattag       240

Db     241 aggaaatacacacgaggcaactcgttgttggacttgttgcaaaatagcaaggtattag       300
Qy     241 taatagcaagcaatattgtagagagcgcaataatgtatacacatatagtagcaag        300

Db     301 taatagcaagcaatattgtagagagcgcaataatgtatacacatatagtagcaag         360
Qy     301 cttacactgtatgaaataaaaaaactatataggtaaagtgaacaacatgcgttatgtaaatc 420

Db     361 cttacactgtatgaaataaaaaaactatataggtaaagtgaacaacatgcgttatgtaaatc 420
Qy     361 tagcagactattggattgaaaagaattcaatactacaagaacaagaaatgactacggaggc 480

Db     421 tagcagactattggattgaaaagaattcaatactacaagaacaagaaatgactacggaggc 480
Qy     421 agcaacacaataactaaacatgttccaaaaatgtcagatatgaaggctccgaacgccatgcac 540

Db     481 agcaacacaataactaaacatgttccaaaaatgtcagatatgaaggctccgaacgccatgcac 540
Qy     481 ggcatgatatgctagttgggccgtttccgtcgggcttaaagataaggaaatctggata 600

Db     541 ggcatgatatgctagttgggccgtttccgtcgggcttaaagataaggaaatctggata 600
Qy     541 tggactaatgatgtctaatttttgttagagcctagaagtgactaactagaaggt 660

Db     601 tggactaatgatgtctaatttttgttagagcctagaagtgactaactagaaggt 660
Qy     601 taattttgttctattttttgttgcaccgactgagccaactcttttgtctagtagttt 720

Db     661 taattttgttctattttttgttgcaccgactgagccaactcttttgtctagtagttt 720
Qy     661 acatttagttactactctcctttcgtctaaaaagtactacatatctcatttttaaatgtc 780

Db     721 acatttagttactactctcctttcgtctaaaaagtactacatatctcatttttaaatgtc 780
Qy     721 ttgctttttgaagagcactatctctttaaatcttgaccaactatataaagtacttctga 840

Db     781 ttgctttttgaagagcactatctctttaaatcttgaccaactatataaagtacttctga 840
Qy     781 tacatgataggtttaataaaatcttattattttttagtctcagtcaaact 900

Db     841 tacatgataggtttaataaaatcttattattttttagtctcagtcaaact 900
Qy     841 taagagctttgatcgcacatagttgtttaaacaaggtcgttgttcatgtcgttct 960

Db     901 taagagctttgatcgcacatagttgtttaaacaaggtcgttgttcatgtcgttct 960
Qy     901
```

```
Db     961 aatatgtcggatagtattccgattcattccgccagaggcgttgctcgttggatattggttag 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     961 agcatcttcaagaaaacccgtaaatccaactccaaaaacgttttgagcctccaacgtcc 1080

Db    1021 agcatcttcaagaaaacccgtaaatccaactccaaaaacgttttgagcctccaacgtcc 1080
Qy    1021 gttcccacgtgtcattctttcctcgcgcgcgcagttctgttgtttggagaggaagcggga 1140

Db    1081 gttcccacgtgtcattctttcctcgcgcgcgcagttctgttgtttggagaggaagcggga 1140
Qy    1081 actagcgctggctgggagattatggccatgcaatcagttttgttcagtcaaatgctttgt 1200

Db    1141 actagcgctggctgggagattatggccatgcaatcagttttgttcagtcaaatgctttgt 1200
Qy    1141 ttttttggccgtcatggtgagtatcttcattctggaaaactctaaataaatattgaatgt 1320

Db    1201 ttttttggccgtcatggtgagtatcttcattctggaaaactctaaataaatattgaatgt 1320
Qy    1201 atatttctttaagtcgtcacgtataggataggaacctaaaaccatgc 1380

Db    1321 atatttctttaagtcgtcacgtataggataggaacctaaaaccatgc 1380
Qy    1321 acatattcatctctttcctcaaaagcaagtctgtttaataatattattctcgttccagctaga 1440

Db    1381 acatattcatctctttcctcaaaagcaagtctgtttaataatattattctcgttccagctaga 1440
Qy    1381 ttatccttagaaattagttctaaaacgcaagtctcggtgatattctcgttgccagctaga 1500

Db    1441 ttatccttagaaattagttctaaaacgcaagtctcggtgatattctcgttgccagctaga 1500
Qy    1441 ccatggtttaaactattacggtgtatacagtgactaccttagtattttaaagtgatcag 1560

Db    1501 ccatggtttaaactattacggtgtatacagtgactaccttagtattttaaagtgatcag 1560
Qy    1501 ttgaaatatgtcgacttattttagtatttattataaagatacagtgttcctgtgccagctaga 1620

Db    1561 ttgaaatatgtcgacttattttagtatttattataaagatacagtgttcctgtgccagctaga 1620
Qy    1561 acatattcatctctttcctcaaaagcaagtctgtattttaataataattatca 1620

Db    1621 catttaaacacaatctcaatcatatattgttgctatactagttttttgctttcagtat 1680
Qy    1621 catttaaacacaatctcaatcatatattgttgctatactagttttttgctttcagtat 1680

Db    1681 gcctaccatcgggatcgctcgtatacaagtgactaactttaaatgaagtcgcga 1740
Qy    1681 gcctaccatcgggatcgctcgtatacaagtgactaactttaaatgaagtcgcga 1740

Db    1741 aaattatataaggttacttagaggacggaagtagtgaagtggtaaacaaatc 1800
Qy    1741 aaattatataaggttacttagaggacggaagtagtgaagtggtaaacaaatc 1800

Db    1801 aactaaaatttaaattccaagtaatattataaattgactgactttcttcagtat 1860
Qy    1801 aactaaaatttaaattccaagtaatattataaattgactgactttcttcagtat 1860

Db    1861 cttataatggttactataggttgaggaaggtagtgaactggtaaacaaatc 1920
Qy    1861 atatttagcagctcgctaaaatctgaattgaaaatttgaagtgaatgct 1920

Db    1921 cttataatggttactataggttgaggaaggtagtgaactggtaaacaaatc 1980
Qy    1921 cttatataatggttactataggttgaggaaggtagtgaactggtaaacaaattc 1980

Db    1981 gaattgatctattcaacttgtagctactcagcaagatgcgaatcaaatccggcg 2040
Qy    1981 gaattgatctattcaacttgtagctactcagcaagatgcgaatcaaatccggcg 2040

Db    2041 gggtggattccgcacgcccacggctgggttcgttcgttctcacgccgttcaatctc 2100
Qy    2041
```

```
Qy 2041 gggtcggattccgccacggccacggggtggggttcggtcgtctcaccgcggtcaatctc
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 ccctccgcgcggcgcaatctcgtcccggtggggacggcggctagctgcccaatctc 2160
Qy 2101 ccctccgcgcggcgcaatctcgtcccggtggggacggcggctagctgcccaatctc
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 caccgacacaaagcgccatgcgtgttccgtacaattggcctccttccccgcgcc 2220
Qy 2161 caccgacacaaagcgccatgcgtgttccgtacaattggcctccttccccgcgcc
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 ttcctcctttcctgcgtgaccgcaacacactgcgctcaccacgtgtacatgtattc 2280
Qy 2221 ttcctcctttcctgcgtgaccgcaacacactgcgctcaccacgtgtacatgtattc
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 tccctagccgaaccgtatcaagtagtcttaggggtggggcgttcggtttaccgaaatttt 2340
Qy 2281 tccctagccgaaccgtatcaagtagtcttaggggtggggcgttcggtttaccgaaatttt
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 cgggttgggtaattcaagtttttttaaattcgggtttttgagaatcaataccgaaattac 2400
Qy 2341 cgggttgggtaattcaagtttttttaaattcgggtttttgagaatcaataccgaaattac
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 aacggatttttcaataccggaattcggggttacccggggttcggggta 2460
Qy 2401 aacggatttttcaataccggaattcggggttacccggggttcggggta
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 ttcccaaactaccggaactaccggggttcggggttgcttcatataaaaacactacacctattaaatt 2520
Qy 2461 ttcccaaactaccggaactaccggggttcggggttgcttcatataaaaacactacacctattaaatt
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 agtataaaatataagtttgaataatgatataatacacatggacacataaaacaaacaatctac 2580
Qy 2521 agtataaaatataagtttgaataatgatataatacacatggacacataaaacaaacaatctac
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 aatcccaagttatgcacacttacacataattatagatgtacaacttaaattattaaagc 2640
Qy 2581 aatcccaagttatgcacacttacacataattatagatgtacaacttaaattattaaagc
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 atgacatgagtacatgacacatgaaagccgggtaattcgggtgtattcgggtacccgattg 2700
Qy 2641 atgacatgagtacatgacacatgaaagccgggtaattcgggtgtattcgggtacccgattg
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 tgatacccgaattaccggaataattcgggtttcaagttgtcaccggaaatccccaa 2760
Qy 2701 tgatacccgaattaccggaataattcgggtttcaagttgtcaccggaaatccccaa
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 acaaaatcgggtttcggtattcgggttcgggttcggttcggtttgggtttcg 2820
Qy 2761 acaaaatcgggtttcggtattcgggttcgggttcggttcggtttgggtttcg
         ||
Db 2821 gg 2822
Qy 2821 gg 2822
```

RESULT 2

CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli, beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within this way are used as primers for
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 diff base substitutions, most of which
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence 204 BP;   21 A;   47 C;   17 G;   11 T;   108 Others;

Query Match   1.7%;   Score 49;   DB 1;   Length 204;
Best Local Similarity   15.4%;
Matches   18;   Conservative   58;   Mismatches   41;   Indels   0;   Gaps   0;

```
Db   69  cycgaaycycdchvgccgymttthyrrmtbnvyrdynrsdaaawycyirsvkydccyn 128
         |:|||  |:.||:.|:||  :|::            |.||:|:|  ::
Qy  2417  cccggaattcgggtaccgggttcgggtattcccaaactaccga 2476
Db   129  achddhyyybbbvynvhnhnncncccbnhhvchnvhbnnhnrnwayvrhdarrddvh 185
Qy  2477  actattgtgttggtccataaaaacacacacactattaaattagtataaaaatat 2533
```

RESULT 3

CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp. The probe may
CC   be useful as an initial screen for mycobacterial infection.
SQ   Sequence 91 BP;   5 A;   17 C;   15 G;   4 T;

Query Match   1.6%;   Score 44;   DB 9;   Length 91;
Best Local Similarity   7.7%;   Pred. No. 3.73e-10;
Matches   4;   Conservative   44;   Mismatches   4;   Indels   0;   Gaps   0;

```
Db   13  vhsyyvvhvsshhvshhvvvvhhvvhvvhhvyhvyysvctca 64
         :::::::::::::::::::::::::::::: ::  :::::||||
Qy 2209  ccctcccgcccctccctctcctctgccgtgacgcaacacactgcgctcca 2260
```

RESULT 4

DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
SS   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993;
PD   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT) BECTON DICKINSON CO.
CC   Shank DD, Spears PA;
PI   WPI 93-378844/48.
DR
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
     samples
PS   Claim 3; Page 14; 23pp; English.
PT   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp. The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match   1.5%; Score 42; DB 9; Length 91;
Best Local Similarity 18.1%; Pred. No. 4.74e-09;
Matches 13; Conservative 44; Mismatches 15; Indels 0; Gaps 0;

Db   9 cggsvhsyyvvhvvvshhhvvhvvhhvsvvvhhvvhvhyhyysvctcaagc 68
         ||| :|||||| :|::::   ::||||| ::::||:::|:|:|::: |||
Cp 1195 cgtcccgccttcctccaacagaactgcgcgcgcgaggaaaatacacgtgggaaccagca 1136

Db   69 tcggcggcgaac 80
         |||| ||||
Cp 1135 tgcgcagaggac 1124

RESULT 5
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DT   08-NOV-1990 (first entry)
DE   Base substituted E.coli beta-galactosidase alpha-fragment.
DE   E.coli beta galactosidase alpha-fragment; base substitution; ss.
KW   E.coli beta-galactosidase alpha-fragment.
OS   Escherichia coli.
FH   Key   Location/Qualifiers
FT   misc_feature   19..69
FT   /*tag= a
FT   /function=multiple cloning site
FT   primer_bind   187..204
FT   /*tag= b
PN   EP-285123-A.
PD   05-MAY-1988;
PD   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
CC   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
PI   WPI 88-279927/40.
DR
PT   Introducing random point mutations into nucleic acds -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure; P; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   reverse transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 diff base substitutions, most of which
CC   occurred singularly in any given mutant.
SQ   Sequence 204 BP; 21 A; 47 C; 17 G; 108 Others;

Query Match   1.5%; Score 42; DB 1; Length 204;
Best Local Similarity 15.1%; Pred. No. 4.74e-09;
Matches 14; Conservative 49; Mismatches 29; Indels 1; Gaps 1;

Db   91 thhyrrmrbnvyrdyrnrsdaaawyccyrrsvykydccynachhddhyvybbbvynvhnhnn 150
         ||:: ::: ||: :::: :::||||: ||: ::::||: :::::||::::|:|::
Cp 2383 ttcttcaaaacccgaaatttaaaaaaacttgaattaccaacccgaaaaattccg-ggtaacc 2325

Db   151 cnccbnnhvchnvhbnnhrnwayyrhdarrdd 183
         | :: ||: : |: :: : :|: :|::||:| |||
Cp 2324 cgaacgcccacccctaagaactactgataacggt 2292

RESULT 6
ID   V44650 standard; DNA; 91 BP.
AC   V44650;
DT   06-OCT-1998 (first entry)
DE   Mammalian DNA replication origin consensus sequence, uniorsconsensus.
KW   DNA replication origin; human; mammal; alphaconsensus; uniorsconsensus;
KW   anti-gene; DNA replication inhibitor; shuttle vector construct creation;
KW   gene therapy; ss.
OS   Mammalia.
PN   WO9827200-A2.
PD   25-JUN-1998.
PD   12-DEC-1997; CA0972.
PR   21-MAY-1997; US-047322.
PR   16-DEC-1996; US-033374.
PA   (UYMC-) UNIV MCGILL.
PA   Cossons NH, Nielsen TO, Price GB, Zannis-Hadjopoulos M;
DR   WPI 98-362770/31.
PI   Human or mammalian origin of replication consensus sequences - for
PT   inhibiting DNA replication, for controlling initiation of
PT   replication, maintaining circular plasmids and in assembly of human
     artificial chromosomes
PS   Claim 1; Page 42; 54pp; English.
CC   This sequence represents a human or mammalian DNA replication origin
CC   consensus sequences of the invention, designated uniorsconsensus.
CC   Administration of the consensus sequence or an anti-gene (comprising a
CC   double stranded copy of the consensus) is used to inhibit DNA replication
CC   in vivo or in vitro. The consensus sequences can also be inserted into an
CC   expression vector, used subsequently for in vitro transfection of
CC   mammalian cells, to control initiation of DNA replication. They can also
CC   be used to maintain circular plasmids that are capable of
CC   semi-conservative replication in proliferating mammalian cells, or
CC   inserted into mammalian or human artificial chromosome vectors for gene
CC   therapy. Particularly, they are used to create shuttle vector constructs
CC   for defining the essential mammalian elements required for maintenance of
CC   chromosomal function. The consensus sequence can be combined with cloned
CC   human telomeres and large centromeric blocks for assembly of human
CC   artificial chromosomes and maintained as bacterial plasmids, circular or
CC   linear, large or small yeast artificial chromosomes (YACs) or as episomal
CC   elements.
SQ   Sequence 91 BP; 15 A; 1 C; 4 G; 7 T;

Query Match   1.28%; Score 33; DB 46; Length 91;
Best Local Similarity 10.8%; Pred. No. 2.82e-04;
Matches 9; Conservative 49; Mismatches 25; Indels 0; Gaps 0;

Db   8 krawrwwkkdavrwwgakrwwkwvwhrassacmdwkaaktwkggwtwarrywkgrkmwwtw 67
         ::: :: |: ::: :::|: ::||: |||::|:|:: :: ::: ::|:: ::::|
Qy 1509 taaactattacggtaaaatagcatgacttagctatttaaaatgatatacagttgaaata 1568

Db   68 kawsdatakwwkdakwkmwrkt 90
         :: ::|:|:::|:||::::::|
Qy 1569 tgtcgacttatttatagttagt 1591

RESULT 7
ID   V44650 standard; DNA; 91 BP.
AC   V44650;
DT   06-OCT-1998 (first entry)
DE   Mammalian DNA replication origin consensus sequence, uniorsconsensus.
KW   DNA replication origin; human; mammal; alphaconsensus; uniorsconsensus;

KW   anti-gene; DNA replication inhibitor; shuttle vector construct creation;
KW   gene therapy; DNA replication; ss.
OS   Mammalia.
PN   WO9827200-A2.
PD   25-JUN-1998.
PF   12-DEC-1997; WO9827200-A2.
PF   21-MAY-1997; US-047322.
PR   16-DEC-1996; US-033374.
PA   (UYMC-) UNIV McGILL.
PI   Cossons NH, Nielsen TO, Price GB, Zannis-Hadjopoulos M;
DR   WPI; 98-362770/31.
DR   R65151-54.
PT   Human or mammalian origin of replication consensus sequences - for
PT   inhibiting DNA replication, for controlling initiation of
PT   replication, maintaining circular plasmids and in assembly of human
PT   artificial chromosomes
PS   Claim 1; Page 42; 54pp; English.
CC   This sequence represents a human or mammalian DNA replication origin
CC   consensus sequences of the invention, designated uniorsconsensus
CC   consensus sequences of the consensus sequence or an anti-gene (comprising a
CC   double stranded copy of the consensus) is used to inhibit DNA replication
CC   in vivo or in vitro. The consensus sequences can also be inserted into an
CC   expression vector, used subsequently for in vitro transfection of
CC   mammalian cells, to control initiation of DNA replication. They can also
CC   be used to maintain circular plasmids that are capable of
CC   semi-conservative replication in proliferating mammalian cells, or
CC   inserted into mammalian or human artificial chromosome vectors for gene
CC   therapy. Particularly, they are used to create shuttle vector constructs
CC   for defining the essential mammalian elements required for maintenance of
CC   chromosomal function. The consensus sequence can be combined with cloned
CC   human telomeres and large centromeric blocks for assembly of human
CC   artificial chromosomes and maintained as bacterial plasmids, circular or
CC   linear, large or small yeast artificial chromosomes (YACs) or as episomal
CC   elements.
SQ   Sequence   91 BP;   15 A;   1 C;   4 G;   7 T;

Query Match
Best Local Similarity 1.2%;   Score 34;   DB 46;   Length 91;
Matches   9;   Conservative 47;   Mismatches 22;   Indels 0;   Gaps 0;

Qy   11 wrwwkkdavwwgakrwwkwwwhrassacmdwkaaktwkggwtwarrywkgrkmwwtwkaw 70
         ::|| :||| ||: || ||||| |: || || |||| |:
Db 157 tgaatcatattgtaatgttttgtgtgaggcgtgtgaattctttttgttacattagcaacacaa
Cp 157

Db   71 sdatakwwwkdakvkmwr 88
         ::||:|| :||| :|: ::|
Cp   97 gtatgtgagtgttgaacaa 80

RESULT   8
ID   Q70467 standard; DNA; 114 BP.
AC   Q70467;
DT   05-APR-1995 (first entry)
DE   Generic DNA sequence to generate a random TSAR peptide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concatenated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key   Location/Qualifiers
FT   misc_feature   55..60
FT   /*tag= a
FT   /note= "this sequence represents 'z'; z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PF   01-FEB-1994; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65153.

PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows: X(NNB)16(TGC)(NNB)12(TGC)(NNB)16(TGC)(NNB)1Y. X
CC   and Y are flanking restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides.The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need for
CC   complex methods of hybridoma formation or in vivo antibody production.
CC   The TSARs are easily characterised and have designed activity allowing
CC   direct and rapid detection in a screening process.
SQ   Sequence   114 BP;   0 A;   2 C;   2 G;   2 T;

Query Match
Best Local Similarity 1.2%;   Score 34;   DB 12;   Length 114;
Matches   3;   Conservative 32;   Mismatches 71;   Indels 0;   Gaps 0;

Qy 637 gccctagccatgctaactagaagagttaatttgtttccatttttgttgcacggactgagc 696
         |:| || | :| | || || ||| | | ||||
Db   3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnnnnnn 62

Db   63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 108
         : :| || :| :|| || :| :|| ::| :|
Qy 697 caacattccttgtctagtagtttacattttagttactactcctt 742

RESULT   9
ID   T76363 standard; DNA; 172 BP.
AC   T76363;
DT   15-SEP-1997 (first entry)
DE   Human interleukin 8 antisense oligonucleotide.
PN   WO9640152-A1.
PD   19-DEC-1996.
PF   06-JUN-1996; U09306.
PF   07-JUN-1995; US-474497.
PA   (UYEC-) UNIV EAST CAROLINA.
PI   Metzger WJ, Nyce JW;
DR   WPI; 97-051871/05.
KW   Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW   chronic obstructive pulmonary disease; bronchitis; ss.
OS   Synthetic.
PT   Treatment of airway diseases such as asthma - by topically applying
PT   adenosine-free antisense oligo:nucleotide to airway epithelium of
PT   subject
PS   Claim 5; Page 36; 71pp; English.
CC   A method for treating airway disease in a subject has been produced,
CC   which involves the topical administration of an essentially adenosine
CC   free antisense oligonucleotide (ON) to the airway epithelium of the
CC   subject. The present sequence is an antisense oligonucleotide specific
CC   for the human interleukin 8, targeted at the initiation codon. The
CC   method can be used to treat airway diseases such as cystic fibrosis,
CC   asthma, chronic obstructive pulmonary disease, bronchitis and other
CC   airway diseases characterised by an inflammatory response. By
CC   eliminating adenosine from the antisense ON, its liberation upon
CC   antisense degradation is prevented, thereby preventing adenosine-
CC   induced bronchoconstriction in patients with hyper-reactive airways.
SQ   Sequence   172 BP;   0 A;   35 C;   42 G;   39 T;

Query Match
Best Local Similarity 31.2%;   Score 33;   DB 32;   Length 172;
Matches 31;   Conservative 35;   Mismatches 33;   Indels 0;   Gaps 0;
Pred. No. 2.82e-04;

```
Db    40 tbbggbtcbcbtttbgbcbttbggbbgbbbbgggbggbgctgctgctgccbgbbbgbbgbggbgccb 99
         |::|:|:| |:|:| |:||:|:|:||::|| |:|:|::|| |:|:|::| |:|
Qy  2049 tccgccacggccacggtgggtcggttcgttcgttctcaccgccggtcaatctccccccgc 2108
```

```
Db   100 bcbgbgbbgctgcbgbbtcbgbggbbggbgctgccbbgbgbgbgccb 138
         |::|  |:| |:|:|:| |:| |::|:::|| |:|
Qy  2109 gcggcgccaattcgtccggtggggacggctagctggccc 2147
```

RESULT 10

```
ID   Q70465 standard; DNA; 114 BP.
AC   Q70465;
DT   05-APR-1995 (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT                   /*tag= a
FT                   /note= "this sequence represents 'z'; z can be a
FT                   sequence of 6, 9 or 12 nucleotides (see
FT                   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994;
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
```

WPI; 94-279739/34.
P-PSDB; R65150 and R65151.

PS Disclosure; Page 35; 255pp; English.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain

CC Q70465 is a generic DNA sequence used to generate a random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concateneated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides.The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP;

Query Match
Best Local Similarity 1.1%;   Score 32;   DB 12;   Length 114;
Matches 3;   Conservative 4;   Mismatches 32;   Indels 76;   Gaps 0;
Pred. No. 9.05e-04;

```
Db     3 bnnbnbnbnnbtgcnnbnbnnbnbnbnnbnbnnnnnnnnn 62
         ::|   ::  ||:| ::
Qy  1109 gcattgtttccaatcgtcctctcgcgcatgctggtcccacgtgtatttcctccgcgccgca 1168
```

RESULT 11

```
ID   Q70468 standard; DNA; 114 BP.
AC   Q70468;
DT   05-APR-1995 (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT                   /*tag= a
FT                   /note= "this sequence represents 'z'; z can be a
FT                   sequence of 6, 9 or 12 nucleotides (see
FT                   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994;
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
```

WPI; 94-279739/34.
P-PSDB; R65154.

PS Disclosure; Page 35; 255pp; English.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain

CC Q70468 is a generic DNA sequence used to generate a random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)16Z(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68,
CC R65151-54. TSARs are concateneated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. These residues
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP;

Query Match
Best Local Similarity 1.1%;   Score 32;   DB 12;   Length 114;
Matches 3;   Conservative 3;   Mismatches 33;   Indels 76;   Gaps 0;
Pred. No. 9.05e-04;

```
Db     3 bnnbnbnbnnbtgcnnbnbnnbnbnbnnbnbnnnnnnnnn 62
         ::|   ::  ||:|
Qy  1112 ttgttcccaatcgtctctcgcgccatgctgttccacgtgttatttcctcgcgccaagtt 1171
```

RESULT 12

```
Db    63 bnnbnbnbnnbnbtgcnnbnnbnbnbnnbnbnnbnbnnbnbnbnb 114
         :::| |:: |
Qy  1169 gtcctgtttggaggaggaaggcgggacgttggcactagcgctggctggagatt 1220
```

```
Db     3 bnnbnbnbnnbtgcnnbnbnnbnbnbnnbnbnnnnnnnnn 62
         :|:|      ||
Qy  1172 ctgttgaggaggaaggcgggacgttggcactagcgctggctggagattatg 1223
```

ID   Q70469 standard; DNA; 114 BP.
AC   Q70469;
DT   07-APR-1995 (first entry)
DE   Generic DNA sequence to generate a random TSAR peptide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key   Location/Qualifiers
FT   misc_feature   55..60
FT   /*tag= a
FT   /note= "this sequence represents 'z'; z can be a
FT   sequence of 6,9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
PS   Disclosure; Page 35; 255pp; English.
CC   Q70469 is a generic DNA sequence used to generate a random TSAR peptide
CC   This generic formula can be represented as follows: X(TGC)(NNB)10-
CC   (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC   sites (X is not the same as Y) that are not specified further. This
CC   sequence generates peptides that are cloverleaf in structure. Other
CC   generic sequences are shown in Q70465-68. Other specific peptides
CC   generated by these generic sequences are shown in R65150-54. TSARs are
CC   concatenated heterofunctional proteins or peptides, comprising at least
CC   two functional regions - a binding domain with affinity for a ligand and
CC   a second effector peptide portion that is chemically or biologically
CC   active.They may further comprise a linker peptide between the 2 domains.
CC   The oligonucleotides are also designed so that the expressed peptide
CC   contains 2 or 4 cysteine residues positioned in, or flanking, the
CC   unpredicted or variant residues. These residues confer some degree of
CC   conformational rigidity to the peptides.The TSARs or compsns. comprising
CC   a TSAR binding domain can be used in vivo to deliver a chemically or
CC   biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC   or enzyme, to the specific target or on the cell. They can also replace
CC   the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC   and therefore circumvent the need for complex methods of hybridoma
CC   formation or in vivo antibody production. The TSARs are easily
CC   characterised and have designed activity allowing direct and rapid
CC   detection in a screening process.
SQ   Sequence 114 BP;

Query Match   1.1%; Score 31; DB 12; Length 114;
Best Local Similarity   3.7%; Pred. No. 2.8e-03;
Matches   4; Conservative   3; Mismatches 72; Indels   0; Gaps   0;

Qy   6 bnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnnnnnbnn 65
            : :: : | || |: | :::| | |:
Db   1142 ttccacgtgtatttcctcgcgcgcagttctgttggaggaggaggcgggacgttggca 1201

Qy   66 btgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbt 112
            :| :   : :   : ||
Db   1202 ctagcgctggctggagattatggccatcgcaatcagcttggccagt 1248

RESULT 13
ID   Q70470 standard; DNA; 114 BP.
AC   Q70470;
DT   10-APR-1995 (first entry)
DE   Generic DNA sequence to generate a random TSAR peptide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.

OS   Synthetic.
FH   Key   Location/Qualifiers
FT   misc_feature   55..60
FT   /*tag= a
FT   /note= "encoded by z (see comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R58378.
PS   Disclosure; Page 36; 255pp; English.
CC   Q70470 is a generic DNA sequence used to generate a random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows: X(NNB)4(CAC)(NNB)8Z(NNB)6(CAC)(NNB)8
CC   -(CAC)2(NNB)Y. X and Y are flanking restriction sites (X is not the same
CC   as Y) that are not specified further. The peptides generated by this and
CC   other generic sequences (Q70471-73) have invariant histidine residues
CC   incorporated into variant sequences. TSARs are concatenated
CC   heterofunctional proteins or peptides, comprising at least two functional
CC   regions - a binding domain with affinity for a ligand and a second
CC   effector peptide portion that is chemically or biologically active.They
CC   may further comprise a linker peptide between the 2 domains.The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed
CC   activity allowing direct and rapid detection in a screening process.
SQ   Sequence 114 BP;

RESULT 14
Query Match   1.1%; Score 31; DB 12; Length 114;
Best Local Similarity   5.8%; Pred. No. 2.8e-03;
Matches   6; Conservative 29; Mismatches 68; Indels   0; Gaps   0;

Qy   3 bnnbnnbcacnnbnnbnnbcacnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnn 62
Db   1136 tgctggttccacgtgtatttcctcgcgcgcagttctgttggaggaggaggcgggacg 1195

Qy   63 bnnbnnbnnbnnbcacnnbnnbnnbnnbnnbnnbnnbnnbnnb 105
Db   1196 tggcactagcgctggctggagattatggccatcgcaatcagcagt 1238

```
RESULT 15
ID   Q70469 standard; DNA; 114 BP.
AC   Q70469;
DT   07-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR peptide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concatenated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FH
FT   misc_feature    55..60
FT                   /*tag=      a
FT                   /note=      "this sequence represents 'z'; z can be a
FT                               sequence of 6,9 or 12 nucleotides (see
FT                               comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; US00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UINC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
```

```
CC   Disclosure; Page 35; 255pp; English.
CC   Q70465 is a generic DNA sequence used to generate a random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides. This generic formula can also be
CC   represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)3Y. X
CC   and Y are flanking restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70465-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active. They may further comprise a linker
CC   peptide between the 2 domains. The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion, on the
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed
CC   activity allowing direct and rapid detection in a screening process.
SQ   Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match            1.1%;  Score 30;  DB 12;  Length 114;
Best Local Similarity  4.9%;
  Matches    5;  Conservative   28;  Mismatches   69;  Indels   0;  Gaps   0;
                          Pred. No. 8.93e-03;

Qy          3 bnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 62
                |||       |||
Db         63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtg 104

Cp        555 ctagccatatcatgccgtcgccgttcgagcccttcatatctgaccattggaacatt 496

Db        495 agttattgttgctgccctcgtccgtcagtcattcttgccttg 454
```

```
PS   Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC   Disclosure; Page 35; 255pp; English.
CC   This generic formula can be represented as follows: X(TGC)(NNB)10-
CC   (TGC)(NNB)6Z(TGC)(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC   sites (X is not the same as Y) that are not specified further. This
CC   sequence generates peptides that are cloverleaf in structure. Other
CC   generic sequences are shown in Q70465-68. Other specific peptides
CC   generated by these generic sequences are shown in R65150-54. TSARs are
CC   concatenated heterofunctional proteins or peptides, comprising at least
CC   two functional regions - a binding domain with affinity for a ligand and
CC   a second effector peptide portion that is chemically or biologically
CC   active. They may further comprise a linker peptide between the 2 domains.
CC   The oligonucleotides are also designed so that the expressed peptide
CC   contains 2 or 4 cysteine residues positioned in, or flanking, the
CC   unpredicted or variant residues. These residues confer some degree of
CC   conformational rigidity to the peptides. The TSARs or compsns. comprising
CC   a TSAR binding domain can be used in vivo to deliver a chemically or
CC   biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC   or enzyme, to the specific target or on the cell. They can also replace
CC   the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC   and therefore circumvent the need for complex methods of hybridoma
CC   formation or in vivo antibody production. The TSARs are easily
CC   characterised and have designed activity allowing direct and rapid
CC   detection in a screening process.
SQ   Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match            1.1%;  Score 30;  DB 12;  Length 114;
Best Local Similarity  4.6%;
  Matches    5;  Conservative   30;  Mismatches   73;  Indels   0;  Gaps   0;
                          Pred. No. 8.93e-03;

Qy          1 tgcnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbtgccagctagccgtccccaccggacgaa 60
                |||                   |||
Db       2177 tgcggcatttgtcggtggagctttggccattggccagctagccgtccccaccggacgaa 2118

Db         61 nnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 60

Cp       2117 ttgcgcgccggaggggagattgacacggcgcggtgagaacgacacgaac 2070
```

Search completed: Tue Dec 28 20:54:38 1999
Job time : 1750 secs.